

Introduction to labeled island particle models and their asymptotic properties

Cécile Ichard¹ and Christelle Vergé^{2,3,4}

¹Météo-France-CNRS, CNRM-GAME UMR 3589, 42 Avenue Coriolis, 31057
Toulouse Cedex 1, France; email : cecile.ichard@meteo.fr

²Onera - The French Aerospace Lab, Chemin de la Hunière et des Joncherettes,
BP 80100, 91123 Palaiseau Cedex, France; email : christelle.verge@onera.fr

³CNES - Centre National d'Etudes Spatiales, 18 avenue Edouard Belin, 31 401
Toulouse Cedex 9, France

⁴CMAP UMR 7641 École Polytechnique CNRS, Route de Saclay, 91128
Palaiseau Cedex France

March 26, 2015

Abstract

Estimation of stochastic processes evolving in a random environment is of crucial importance for example to predict aircraft trajectories evolving in an unknown atmosphere. For fixed parameter, interacting particle systems are a convenient way to approximate such stochastic process. But the second level of uncertainty provided by the environment parameters leads us to also consider interacting particles on the parameter space. This novel algorithm is described in this paper. It allows to approximate both a random environment and a stochastic process evolving in this environment, given noisy observations of the process. It is a sequential algorithm that generalizes island particle models including a parameter. It is referred by us as labeled island particle algorithm. We prove the convergence of the labeled island particle algorithm and we establish \mathbb{L}^p bound as well as time uniform \mathbb{L}^p bound for the asymptotic error introduced by this double level of approximation. Finally, we illustrate these results on a filtering problem where one learns a dynamical parameter through noisy observations of a stochastic process influenced by the parameter.

Introduction

This paper deals with the estimation of stochastic processes whose evolution is influenced by a random environment. This question is at stake in different areas. In economy, when one wants to estimate the option price with an unknown volatility [1] using the Black-Scholes model, one can consider that the option price has its evolution influenced by an unknown environment, the market volatility. In biology, when one wants to estimate the number of bacteria whereas the environment factors are unknown [2], one can model the evolution of the bacteria number as a stochastic process whose evolution is influenced by unknown external factors. In air traffic management, this modelization can also be used when one wants to predict aircraft trajectories evolving in an unknown atmosphere. Indeed if pilot intents and some aircraft parameters are

not known, actual wind and temperature evolve locally and are not perfectly known neither. Those atmospheric parameters which appear in the dynamic equations of the aircraft have a great importance to predict the future position of the aircraft. They are thus both uncertain. Therefore, in order to improve the trajectory prediction, one has to learn aircraft parameters but also atmospheric ones. It has been shown in [3] that it can be done using mode-S radar observations and this specific model.

When the stochastic process evolving in the random media is Gaussian and its evolution is linear, the double estimation can be made using interacting Kalman filters (IKF) [4, 5]. However when the dynamics are non-linear, as for aircraft dynamics, an analytic resolution is not possible. A method based on interacting particle systems, which takes into account the randomness due to the environment and also the randomness coming from the process itself, was proposed by Del Moral in [4]. This idealized algorithm would be a sequential Monte Carlo (SMC) algorithm on the couple defined by the random environment and the conditional law of the process evolving in this random environment given the history of the environment. However, the calculation of the previous conditional law is not tractable in practice when the dynamics are non linear. Therefore another approximation level is necessary in order to estimate this conditional law. We propose in this paper to use interacting systems of interacting particles. These interacting systems can be viewed as a two-level interacting particle system. The top level particles are composed of an environment proposition and an empirical measure which gives an approximation of the process law evolving in the proposed environment. The empirical measure is obtained by the second level of interacting particles. This nested structure was also presented in [6] for mean field processes.

This algorithm can be seen as a generalization of interacting island particle models where each island is associated with a random parameter. Those island particle models have been introduced in [7] and their statistical properties studied in [8], but without parameters. The first paper deals with the parallelization of interacting particle systems, the second one studies the asymptotic properties of the ensuing estimator. Concerning filtering problems, Chopin et al. in [9] introduced a kind of island particle models where each island is identified by a parameter proposition. They proposed an algorithm called SMC² which is a practical version of the idealized iterated batch importance sampling (IBIS) algorithm introduced by Chopin in [10] for exploring a sequence of parameter posterior distributions. The considered parameter did not have any proper dynamic whereas in the present paper the stochastic process evolution scheme depends on a dynamic parameter. Moreover, in the SMC² algorithm, islands of particles grow continuously with time as particles ancestral lines are required to estimate the likelihood increments, and by their product to estimate the total likelihood. The algorithm introduced by Crisan et al. in [11] is a different version of the SMC² which allows also the estimation of fixed parameters of a state-space dynamic system using sequential Monte Carlo methods. However, unlike the SMC² method, the proposed algorithm by Crisan et al. operates in a purely sequential and recursive manner. In particular, the scheme for the rejuvenation of the particles in the parameter space is simpler, given that it does not need the simulation of the auxiliary particle filter from initial time to evaluate the likelihood. Therefore the algorithm we propose in this paper is similar to the algorithm of [11] in the sense that it is sequential in time and structured as a nested interacting particle filters, but different as it deals with dynamic parameters.

In this article, we present a novel algorithm for estimating both a random environment and a process whose evolution depends on this environment, and study the asymptotic properties of the ensuing estimators. This study is of great importance to justify the convergence of this algorithm and also a challenging issue as it deals with error in distribution space. Therefore as a first step we establish \mathbb{L}^p bound for the asymptotic error introduced by this double level of approximation at every time step. As a matter of fact, the shape of the bound was predicted by Baehr in his thesis [6]. Then we obtain a time uniform \mathbb{L}^p bound for the error. From there we

deduce the almost sure convergence of the estimator towards the target measure. Afterwards, we compare the labeled island particle algorithm and interacting Kalman filters (IKF) on a filtering example dealing with the evolution of a mobile on a random media. In particular, it appears that the labeled island particle algorithm gives a better estimate of the position and the speed of the mobile than IKF. Finally, the labeled island particle algorithm is applied to another filtering problem where one learns a dynamical parameter through observations of a stochastic process influenced by the parameter. The theoretical results of this paper are illustrated on this example.

Formalization of the problem through Feynman-Kac measures is given in Section 1, then the labeled island particle algorithm is described in Section 2. \mathbb{L}^p bounds of this algorithm are established in Section 3. Finally, convergence of the labeled island particle algorithm and some results proved in Section 4 are illustrated in Section 5 on two filtering examples.

1 Feynman-Kac models in random media

In this section, we first present an example which motivates our study, and then we introduce notations and models.

1.1 Example of process evolution in random media

In this article, one always consider stochastic processes whose evolution are influenced by their surrounding environment. When the environment is unknown, one can be interested in estimating both the environment and the law of the stochastic process itself using observations of the last one. Take a really simple example : a mobile evolving in \mathbb{R}^2 whose dynamics is influenced by an unknown exterior force. This problem can be modeled by the following system of equations

$$\begin{cases} X_{n+1} &= X_n + V_n \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \Delta t + \Theta_{n+1} \Delta t + B_n^X \\ V_{n+1} &= V_n + B_n^V, \end{cases} \quad (1)$$

where X_{n+1} denotes the position of the mobile which depends on Θ_{n+1} and B_n^X a Gaussian noise. The proper speed V_n of the mobile, is known up to a Gaussian white noise B_n^V . α is the course track parameter of the mobile. The vector Θ_{n+1} is random and represents the unknown force acting on the position of the mobile.

We are interested in the estimation of the state of the mobile, which depends on the parameter Θ_{n+1} . We thus need to learn both the force, the speed and the position of the mobile. Consider now that noisy observations Y_n of the mobile's state are available. One has to estimate the quantity $\mathbb{E}[(X_0, V_0, \Theta_0), \dots, (X_n, V_n, \Theta_n) | Y_0, \dots, Y_n]$. Therefore, we need to use a model which can tackle this issue. To this end, the formalism of Feynman-Kac models in random media is well adapted. In Section 1.3, we recall the definitions attached to this model and some important results. For a more detailed review see [4].

1.2 Notations

Let us define some notations used in this paper. For $(m, n) \in \mathbb{Z}^2$ such that $m \leq n$ we denote $\llbracket m, n \rrbracket \triangleq \{m, m+1, \dots, n\} \subset \mathbb{Z}$. We will use the vector notation $a_{m:n} \triangleq (a_m, \dots, a_n)$. Moreover, \mathbb{R}_+ and \mathbb{R}_+^* denote the sets of nonnegative and positive real numbers respectively, and \mathbb{N}^* the set of positive integers.

$N(\mu, \Sigma)$ denotes a multivariate Gaussian distribution with mean μ and covariance matrix Σ .

In the sequel we assume that all random variables are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For some given measurable space (E, \mathcal{E}) we denote by $M(E)$ and $\mathcal{P}(E) \subset M(E)$ the set

of measures and probability measures on (E, \mathcal{E}) , respectively. In addition, we denote by $F(E)$ the set of real-valued measurable functions on (E, \mathcal{E}) and by $\mathcal{B}_b(E) \subset F(E)$ the set of bounded such functions. For any $\nu \in M(E)$ and $f \in F(E)$ we denote by $\nu f \triangleq \int f(x) \nu(dx)$ the Lebesgue integral of f under ν whenever this is well-defined. Now, given also some other (Y, \mathcal{Y}) measurable space, an *unnormalized transition kernel* K from (E, \mathcal{E}) to (Y, \mathcal{Y}) is a mapping from $E \times Y$ to \mathbb{R} such that for all $A \in \mathcal{Y}$, $x \mapsto K(x, A)$ is a nonnegative measurable function on E and for all $x \in E$, $A \mapsto K(x, A)$ is a measure on (Y, \mathcal{Y}) . If $K(x, Y) = 1$ for all $x \in E$, then K is called a *transition kernel* (or simply a *kernel*). The kernel K induces two integral operators, one acting on functions and the other on measures. More specifically, let $f \in F(E)$ and $\nu \in M(E)$ and define the measurable function

$$Kf : E \ni x \mapsto \int f(y) K(x, dy),$$

and the measure

$$\nu K : Y \ni A \mapsto \int K(x, A) \nu(dx),$$

whenever these quantities are well-defined. Finally, let K be as above and let L be another unnormalized transition kernels from (Y, \mathcal{Y}) to some third measurable space (Z, \mathcal{Z}) ; then we define the *product* of K and L as the unnormalized transition kernel

$$KL : E \times Z \ni (x, A) \mapsto \int K(x, dy) L(y, A),$$

from (E, \mathcal{E}) to (Z, \mathcal{Z}) whenever this is well-defined.

1.3 Introduction of Feynman-Kac models

Let Θ_n be a random process on E_n^Θ which influences the evolution of another random process X_n on E_n^X . In order to avoid any confusion, all the quantities which refer to the random process Θ_n (resp. X_n) may be identified by the exponent Θ (resp. X). Let the couple $(\Theta_n, X_n)_{n \in \mathbb{N}}$ be a $E_n \triangleq (E_n^\Theta, E_n^X)$ -valued Markov chain of elementary transition matrix T_n from E_{n-1} to E_n defined by

$$T_n((\theta_{n-1}, x_{n-1}), d(\theta_n, x_n)) \triangleq M_n^\Theta(\theta_{n-1}, d\theta_n) M_{\theta_{n-1}, n}^X(x_{n-1}, dx_n),$$

where M_n^Θ and $M_{\theta_{n-1}, n}^X$ are the transition kernels of the Θ_n and X_n processes from E_{n-1}^Θ to E_n^Θ and from E_{n-1}^X to E_n^X respectively.

Its initial distribution is given by

$$\eta_0(d(\theta_0, x_0)) \triangleq \eta_0^\Theta(d\theta_0) \eta_{\theta_0}^X(dx_0),$$

with $\eta_0^\Theta \in \mathcal{P}(E_0^\Theta)$ and $\eta_{\theta_0}^X \in \mathcal{P}(E_0^X)$, denoting respectively the initial distributions of Θ_0 and X_0 given $\Theta_0 = \theta_0$.

Let $(G_n)_{n \in \mathbb{N}}$ be a collection of bounded measurable functions from E_n to $[0, \infty[$. We define the Feynman-Kac measure associated to the couple (G_n, T_n) with initial distribution η_0 by

$$\begin{aligned} & \mathbb{Q}_{\eta_0, n}(d((\theta_0, x_0), \dots, (\theta_n, x_n))) \\ & \triangleq \frac{1}{Z_n} \left\{ \prod_{p=0}^{n-1} G_p(\theta_p, x_p) \right\} \mathbb{P}_{\eta_0^\Theta, n}^\Theta(d(\theta_0, \dots, \theta_n)) \mathbb{P}_{\theta_0, n}^X(d(x_0, \dots, x_n)), \end{aligned} \quad (2)$$

with the normalizing constant Z_n , given by

$$Z_n \triangleq \mathbb{E}_{\eta_0} \left[\prod_{p=0}^{n-1} G_p(\Theta_p, X_p) \right] > 0,$$

and the two path probabilities

$$\mathbb{P}_{\eta_0^\Theta, n}^\Theta(d(\theta_0, \dots, \theta_n)) \triangleq \eta_0^\Theta(d\theta_0) M_1^\Theta(\theta_0, d\theta_1) \dots M_n^\Theta(\theta_{n-1}, d\theta_n),$$

and

$$\mathbb{P}_{\theta_{0:n}, n}^X(d(x_0, \dots, x_n)) \triangleq \eta_{\theta_0}^X(dx_0) M_{\theta_1, 1}^X(x_0, dx_1) \dots M_{\theta_n, n}^X(x_{n-1}, dx_n).$$

As one may have noticed, given $\Theta_{0:n} = \theta_{0:n}$, the sequence X_n is also a Markov chain of transition kernels $(M_{\theta_n, n}^X)_{n \in \mathbb{N}^*}$ and initial distribution $\eta_{\theta_0}^X$. Then one can associate to it another Feynman-Kac path measure which is called quenched.

Definition 1.1. *The quenched Feynman-Kac path measure associated to the realization $\Theta_{0:n} = \theta_{0:n}$ is defined by*

$$\mathbb{Q}_{\theta_{0:n}, n}^X(d(x_0, \dots, x_n)) \triangleq \frac{1}{Z_{\theta_{0:n}, n}^X} \left\{ \prod_{p=0}^{n-1} G_p(\theta_p, x_p) \right\} \mathbb{P}_{\theta_{0:n}, n}^X(d(x_0, \dots, x_n)),$$

where the quenched normalizing constant $Z_{\theta_{0:n}, n}^X$ is given by

$$Z_{\theta_{0:n}, n}^X \triangleq \mathbb{E}_{\theta_{0:n}} \left[\prod_{p=0}^{n-1} G_p(\theta_p, X_p) \right] > 0.$$

In the rest of the paper the quenched potential functions are denoted by $G_{\theta_p, p}$ and defined as

$$G_{\theta_p, p} : x_p \in E_p^X \mapsto G_{\theta_p, p}(x_p) \triangleq G_p(\theta_p, x_p). \quad (3)$$

To get further into the dynamic, one can define the time marginal of the quenched Feynman-Kac measure also called the quenched Feynman-Kac distribution.

Definition 1.2. *For every realization $\Theta_{0:n} = \theta_{0:n}$, the quenched Feynman-Kac distribution flow $\eta_{\theta_{0:n}, n}^X$ on E_n^X is defined for every $f_n \in \mathcal{B}_b(E_n^X)$ by*

$$\eta_{\theta_{0:n}, n}^X(f_n) \triangleq \gamma_{\theta_{0:n}, n}^X(f_n) / \gamma_{\theta_{0:n}, n}^X(1)$$

with $\gamma_{\theta_{0:n}, n}^X(f_n) \triangleq \mathbb{E}_{\theta_{0:n}} \left[f_n(X_n) \prod_{p=0}^{n-1} G_{\theta_p, p}(X_p) \right]$.

The distribution of X_n depends on the trajectory $\theta_{0:n}$ which is emphasized by denoting the unnormalized quenched Feynman-Kac distribution by $\gamma_{\theta_{0:n}, n}^X$. An important result taken from [4], Proposition 2.6.2] is recalled below.

Proposition 1.1. *The quenched distribution sequence $(\eta_{\theta_{0:n}, n}^X)_{n \in \mathbb{N}}$ satisfies the non linear equation :*

$$\eta_{\theta_{0:n+1}, n+1}^X = \Psi_{\theta_n, n}^X(\eta_{\theta_{0:n}, n}^X) M_{\theta_{n+1}, n+1}^X, \quad (4)$$

where the mapping $\Psi_{\theta_n, n}^X : \mathcal{P}(E_n^X) \rightarrow \mathcal{P}(E_{n+1}^X)$ is given by

$$\Psi_{\theta_n, n}^X(\eta_{\theta_{0:n}, n}^X)(dx_n) \triangleq \frac{1}{\eta_{\theta_{0:n}, n}^X(G_{\theta_n, n})} G_{\theta_n, n}(x_n) \eta_{\theta_{0:n}, n}^X(dx_n). \quad (5)$$

Defining the mapping Φ_{n+1}^X by

$$\begin{aligned} \Phi_{n+1}^X : (E_n^\Theta \times E_{n+1}^\Theta) \times \mathcal{P}(E_n^X) &\rightarrow \mathcal{P}(E_{n+1}^X) \\ ((\theta_n, \theta_{n+1}), \eta_{\theta_{0:n}, n}^X) &\mapsto \Psi_{\theta_n, n}^X(\eta_{\theta_{0:n}, n}^X) M_{\theta_{n+1}, n+1}^X \end{aligned} \quad (6)$$

The non linear recursion (4) can be reformulated as

$$\eta_{\theta_{0:n+1},n+1}^X = \Phi_{n+1}^X((\theta_n, \theta_{n+1}), \eta_{\theta_{0:n},n}^X). \quad (7)$$

Remind that, for a fixed value $\theta_{0:n}$ of the random process $\Theta_{0:n}$, the probability measures $(\eta_{\theta_{0:n},n}^X)_{n \in \mathbb{N}}$ can be approximated recursively thanks to an interacting particle system which evolves successively according to selection step with potentials $G_{\theta_n,n}$ defined in (3) and transition kernels $M_{\theta_n,n}^X$. See [4] for further details. Now, consider that the random environment $\Theta_{0:n}$, where the stochastic process X_n evolves, is not known. Then we focus our interest on the estimation of the couple

$$\bar{X}_n \triangleq (\Theta_n, \eta_{\Theta_{0:n},n}^X) \in \bar{\mathbb{E}}_n \triangleq (\mathbb{E}_n^\Theta \times \mathcal{P}(\mathbb{E}_n^X)), \quad (8)$$

made up of the environment and the law of the process evolving in this environment. The tricky part will be to deal with the probability measure space. First, notice that, as it has been shown in [4], the pair process is a Markov chain.

Proposition 1.2 ([4], Proposition 2.6.3). *\bar{X}_n is a Markov chain with transition kernel \bar{M}_n defined for every function $\bar{f}_n \in \mathcal{B}_b(\bar{\mathbb{E}}_n)$ and $(u, \eta) \in \bar{\mathbb{E}}_n$ by*

$$\bar{M}_n(\bar{f}_n)(u, \eta) \triangleq \int_{\mathbb{E}_n^\Theta} M_n^\Theta(u, dv) \bar{f}_n(v, \Phi_n^X((u, v), \eta))$$

and with initial distribution $\bar{\eta}_0 \in \mathcal{P}(\bar{\mathbb{E}}_0)$ defined by

$$\bar{\eta}_0(d(u, \nu)) \triangleq \eta_0^\Theta(du) \delta_{\eta_{\theta_0}^X}(d\nu).$$

To this Markov chain, one may associate the Feynman-Kac distribution flow $\bar{\eta}_n$ defined for every $\bar{f}_n \in \mathcal{B}_b(\bar{\mathbb{E}}_n)$ by

$$\bar{\eta}_n(\bar{f}_n) \triangleq \bar{\gamma}_n(\bar{f}_n)/\bar{\gamma}_n(\mathbf{1}) \quad (9)$$

where

$$\bar{\gamma}_n(\bar{f}_n) \triangleq \mathbb{E}_{\bar{\eta}_0} \left[\bar{f}_n(\bar{X}_n) \prod_{p=0}^{n-1} \bar{G}_p(\bar{X}_p) \right],$$

and the functions \bar{G}_p are non negative functions defined as follows :

$$\begin{aligned} \bar{G}_p : \bar{\mathbb{E}}_p &\rightarrow [0, \infty[\\ (u, \eta) &\mapsto \bar{G}_p(u, \eta) = \int_{\mathbb{E}_p^X} G_p(u, x) \eta(dx) = \int_{\mathbb{E}_p^X} G_{u,p}(x) \eta(dx) = \eta(G_{u,p}). \end{aligned} \quad (10)$$

Proposition 1.3 ([4], p. 86). *For all $n \in \mathbb{N}$, the sequence $\bar{\eta}_n$ satisfies the following non linear recursive equation :*

$$\bar{\eta}_{n+1} = \bar{\Psi}_n(\bar{\eta}_n) \bar{M}_{n+1} = \bar{\Phi}_{n+1}(\bar{\eta}_n), \quad (11)$$

where for every $\mu \in \mathcal{P}(\bar{\mathbb{E}}_n)$, the application $\bar{\Psi}_n : \mathcal{P}(\bar{\mathbb{E}}_n) \rightarrow \mathcal{P}(\bar{\mathbb{E}}_n)$, is defined by

$$\bar{\Psi}_n(\mu)(\bar{f}_n) = \mu(\bar{G}_n \bar{f}_n)/\mu(\bar{G}_n), \quad (\forall \bar{f}_n \in \mathcal{B}_b(\bar{\mathbb{E}}_n)), \quad (12)$$

and the operator $\bar{\Phi}_n$ is defined by

$$\begin{aligned} \bar{\Phi}_{n+1} : \mathcal{P}(\bar{\mathbb{E}}_n) &\rightarrow \mathcal{P}(\bar{\mathbb{E}}_{n+1}) \\ \mu &\mapsto \bar{\Psi}_n(\mu) \bar{M}_{n+1}. \end{aligned}$$

In the non linear case, (11) cannot be solved analytically. Therefore, in the next section, we introduce an interacting particle system to approximate recursively the sequence of Feynman-Kac probability measures $(\bar{\eta}_n)_{n \in \mathbb{N}}$.

2 Algorithm derivation

This section is about the algorithm associated with the Feynman-Kac distribution flow $\bar{\eta}_n$ defined in (9). One considers the process \bar{X}_n associated with the pair (\bar{G}_n, \bar{M}_n) , where the transition kernel \bar{M}_n is defined in Proposition 1.2 and the potential function \bar{G}_n is defined in (10).

2.1 Idealized interacting particle model

Let N_1 be some positive integer. A N_1 -interacting particle system associated with the sequence $((\bar{G}_n, \bar{M}_n))_{n \in \mathbb{N}}$ and the initial distribution $\bar{\eta}_0$, is a sequence of non-homogeneous Markov chain, denoted by $\bar{X}_n^{[N_1]}$, taking value in the product space $\bar{\mathcal{E}}_n^{N_1}$,

$$\bar{X}_n^{[N_1]} \triangleq (\bar{X}_n^i)_{i=1}^{N_1} = (\bar{X}_n^1, \dots, \bar{X}_n^{N_1}) \in \underbrace{\bar{\mathcal{E}}_n \times \dots \times \bar{\mathcal{E}}_n}_{N_1 \text{ times}}.$$

The initial state of the Markov chain $\bar{X}_0^{[N_1]}$ consists in N_1 independent random variables with common distribution $\bar{\eta}_0$. The interacting particle system $(\bar{X}_n^i)_{i=1}^{N_1}$ explores the state space $\bar{\mathcal{E}}_n$ and with the dynamic given to it, empirically samples the law $\bar{\eta}_n$. Each particle i of the system consists in a random variable $\bar{X}_n^i = (\theta_n^i, \eta_{\theta_n^i, n}^X) \in \bar{\mathcal{E}}_n$. Therefore, the empirical process $\bar{\eta}_n^{N_1}$ is defined by

$$\bar{\eta}_n^{N_1} \triangleq \frac{1}{N_1} \sum_{i=1}^{N_1} \delta_{\bar{X}_n^i}. \quad (13)$$

The elementary transition of the Markov chain $\bar{X}_n^{[N_1]}$ from $\bar{\mathcal{E}}_n^{N_1}$ to $\bar{\mathcal{E}}_{n+1}^{N_1}$ is given for any $\bar{x}_n^{[N_1]} \triangleq (\bar{x}_n^1, \dots, \bar{x}_n^{N_1}) \in \bar{\mathcal{E}}_n^{N_1}$ by

$$\begin{aligned} \mathbb{P}_{\bar{\eta}_0}^{N_1} \left(\bar{X}_{n+1}^{[N_1]} \in d\bar{x}_{n+1}^{[N_1]} \mid \bar{X}_n^{[N_1]} \right) &\triangleq \prod_{i=1}^{N_1} \bar{\Phi}_{n+1}(\bar{\eta}_n^{N_1})(d\bar{x}_{n+1}^i) \\ &= \prod_{i=1}^{N_1} \sum_{j=1}^{N_1} \frac{\bar{G}_n(\bar{X}_n^j)}{\sum_{k=1}^{N_1} \bar{G}_n(\bar{X}_n^k)} \bar{M}_{n+1}(\bar{X}_n^j, d\bar{x}_{n+1}^i), \text{ thanks to (11).} \end{aligned}$$

Thus, the evolution of the particle swarm consists in two steps : a selection and a mutation. In the selection step, the particles $(\bar{X}_n^i)_{i=1}^{N_1}$ are selected multinomially with probability proportional to their potentials $(\bar{G}_n(\bar{X}_n^i))_{i=1}^{N_1}$. Selected particles are identified with a hat on Figure 1. Then the mutation step is performed independently using the kernel \bar{M}_{n+1} . The evolution scheme of the particles is illustrated on Figure 1. Using this algorithm one can empirically sample the measure $\bar{\eta}_n$ at each time step n . Several results are available to qualify the subsequent estimator. However, as one may have noticed, for each θ_n^i the measure $\eta_{\theta_n^i, n}^X$ corresponds to the quenched distributions defined in (1.2). That means that one should have the exact quenched measure associated with the parameter realization $\theta_{0:n}^i$ to use that standard particle algorithm. This can happen in two special cases.

Firstly, one special case is when the transition kernel $M_{\theta_n, n}^X$ is Gaussian and the initial distribution $\eta_{\theta_0}^X$ is Gaussian. Indeed, it turns out that this particle algorithm corresponds to the interacting Kalman filters (IKF) (see [4], [5]). That is a N_1 -interacting particle model which is composed of N_1 particles where the measure value part are Gaussian distributions. In other words, for each particle θ_n^i , one iterative step of the Kalman filter is run to update the measure,

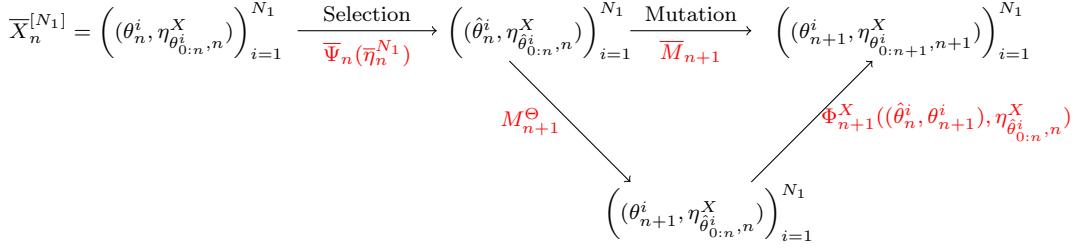


Figure 1: Evolution scheme of the interacting particle system for exact measures.

i.e. one prediction step and one correction step. Those filters are then competing through the selection step using the transformation $\bar{\Psi}_n$ defined in (12). For example let us consider the case where Θ_n is a E_n^Θ -process with initial distribution η_0^Θ and elementary transition kernel M_n^Θ . For a realization $\theta_{0:n}$ of $\Theta_{0:n}$, consider that (X_n, Y_n) is a \mathbb{R}^{p+q} -Markov chain, for positive integers (p, q) , defined through the linear Gaussian system :

$$\begin{cases} X_n &= A_{\theta_n, n} X_{n-1} + a_{\theta_n, n} + B_{\theta_n, n} \varepsilon_n^X, & n \geq 1 \\ Y_n &= C_{\theta_n, n} X_n + c_{\theta_n, n} + D_{\theta_n, n} \varepsilon_n^Y, & n \geq 0. \end{cases}$$

$(A_{\theta_n, n}, B_{\theta_n, n}, C_{\theta_n, n}, D_{\theta_n, n})$ and $(a_{\theta_n, n}, c_{\theta_n, n})$ are respectively matrices and deterministic vectors of appropriate dimension which may depend on a parameter θ_n . The sequences ε_n^X and ε_n^Y are two independent white noises, independent from the initial condition X_0 . There are Gaussian random variables whose mean and variance are given by

$$X_0 \sim N(m_{\theta_0, 0}, \Sigma_{\theta_0, 0}), \quad \varepsilon_n^X \sim N(0, \Sigma_n^X), \quad \text{and } \varepsilon_n^Y \sim N(0, \Sigma_n^Y).$$

In this framework, $\eta_{\theta_{0:n}, n}^X$ corresponds to the conditional law of X_n given the observations $Y_{0:n-1} = y_{0:n-1}$ and the history of the parameter $\theta_{0:n}$, also called optimal predictor. One wants to estimate recursively the law of the couple $(\Theta_n, \eta_{\theta_{0:n}, n}^X)$ using observations $Y_{0:n-1} = y_{0:n-1}$. For that purpose, one needs to introduce the optimal filtering which is the conditional law of X_n given the observations $Y_{0:n} = y_{0:n}$ and the history of the parameter $\theta_{0:n}$. It turns out that these previous distributions are Gaussian respectively denoted by $\eta_{\theta_{0:n}, n}^X = N(m_{\theta_n, n}, \Sigma_{\theta_n, n})$ and $N(\hat{m}_{\theta_n, n}, \hat{\Sigma}_{\theta_n, n})$. Thus,

$$\begin{aligned} \hat{m}_{\theta_n, n} &= \mathbb{E}_{\theta_{0:n}}[X_n | Y_{0:n}] \\ \hat{\Sigma}_{\theta_n, n} &= \mathbb{E}_{\theta_{0:n}}[(X_n - \hat{m}_{\theta_n, n})(X_n - \hat{m}_{\theta_n, n})^T] \\ m_{\theta_{n+1}, n+1} &= \mathbb{E}_{\theta_{0:n+1}}[X_{n+1} | Y_{0:n}] \\ \Sigma_{\theta_{n+1}, n+1} &= \mathbb{E}_{\theta_{0:n+1}}[(X_{n+1} - m_{\theta_{n+1}, n+1})(X_{n+1} - m_{\theta_{n+1}, n+1})^T]. \end{aligned}$$

Moreover, the mapping Φ_{n+1}^X defined in (6) which is used to update the measure valued part $\eta_{\theta_{0:n}, n}^X$ corresponds to a complete step of the Kalman filter evolution between predictors. This means that $\Phi_{n+1}^X((\theta_n, \theta_{n+1}), N(m_{\theta_n, n}, \Sigma_{\theta_n, n}))$ is also a Gaussian distribution whose mean and covariance matrix are obtained recursively through two steps:

$$N(m_{\theta_n, n}, \Sigma_{\theta_n, n}) \xrightarrow{\text{Correction}} N(\hat{m}_{\theta_n, n}, \hat{\Sigma}_{\theta_n, n}) \xrightarrow{\text{Prediction}} N(m_{\theta_{n+1}, n+1}, \Sigma_{\theta_n, n}).$$

The first one is a correction step which is given by

$$\begin{cases} \hat{m}_{\theta_n, n} &= m_{\theta_n, n} + K_{\theta_n, n}(Y_n - (C_{\theta_n, n}m_{\theta_n, n} + c_{\theta_n, n})) \\ \hat{\Sigma}_{\theta_n, n} &= (I - K_{\theta_n, n}C_{\theta_n, n})\Sigma_{\theta_n, n} \end{cases}$$

where I is the identity matrix and $K_{\theta_n, n}$ is the classical gain matrix

$$K_{\theta_n, n} \triangleq \Sigma_{\theta_n, n}(C_{\theta_n, n})^T \left(C_{\theta_n, n}\Sigma_{\theta_n, n}(C_{\theta_n, n})^T + D_{\theta_n, n}\Sigma_n^Y(D_{\theta_n, n})^T \right)^{-1}.$$

The second step is the predicting step :

$$\begin{cases} m_{\theta_{n+1}, n+1} &= A_{\theta_{n+1}, n+1}\hat{m}_{\theta_n, n} + a_{\theta_{n+1}, n+1} \\ \Sigma_{\theta_{n+1}, n+1} &= A_{\theta_{n+1}, n+1}\hat{\Sigma}_{\theta_n, n}(A_{\theta_{n+1}, n+1})^T + B_{\theta_{n+1}, n+1}\Sigma_{n+1}^X(B_{\theta_{n+1}, n+1})^T. \end{cases}$$

Then all the Kalman filters attached to each realization θ_{n+1}^i for $i \in \llbracket 1, N_1 \rrbracket$ interact through their potential $\bar{G}_{n+1}(\theta_{n+1}^i, \eta_{\theta_{0:n+1}^i, n+1}^X)$ defined in (10) by

$$\begin{aligned} \bar{G}_{n+1}(\theta_{n+1}^i, \eta_{\theta_{0:n+1}^i, n+1}^X) &= \eta_{\theta_{0:n+1}^i, n+1}^X(G_{\theta_{n+1}^i, n+1}) \\ &= \mathbb{N}(m_{\theta_{n+1}^i, n+1}, \Sigma_{\theta_{n+1}^i, n+1})(G_{\theta_{n+1}^i, n+1}) \end{aligned}$$

where $G_{\theta_{n+1}^i, n+1}$ is the likelihood function defined for every $x_{n+1} \in \mathbb{E}_{n+1}^X$ by

$$G_{\theta_{n+1}^i, n+1}(x_{n+1}) = \frac{d\mathbb{N}(C_{\theta_{n+1}^i, n+1}x_{n+1}, \Sigma_{n+1}^Y)}{d\mathbb{N}(0, \Sigma_{n+1}^Y)}.$$

One finally ends up with the following expression:

$$\begin{aligned} \bar{G}_{n+1}(\theta_{n+1}^i, \eta_{\theta_{0:n+1}^i, n+1}^X) &= \frac{d\mathbb{N}(C_{\theta_{n+1}^i, n+1}m_{\theta_{n+1}^i, n+1}, C_{\theta_{n+1}^i, n+1}\Sigma_{\theta_{n+1}^i, n+1}(C_{\theta_{n+1}^i, n+1})^T + \Sigma_{n+1}^Y)}{d\mathbb{N}(0, \Sigma_{n+1}^Y)}. \quad (14) \end{aligned}$$

See [4] for further details. The interacting Kalman filter for this general example is given by Algorithm 1.

Data: $\bar{\eta}_0$, $(\bar{M}_p)_{p=0}^n$, $(\bar{\Psi}_p)_{p=0}^n$, $m_{\theta_0^i,0}$ and $\Sigma_{\theta_0^i,0}$

Result: Interacting Kalman approximation of $\bar{\eta}_n$

```

/* Initialization */
```

for $i \leftarrow 1$ **to** N_1 **do**

Sample $\tilde{X}_0^i = (\theta_0^i, \eta_{\theta_0^i,0}^X) \sim \tilde{\eta}_0$, i.e. $\theta_0^i \stackrel{i.i.d.}{\sim} \eta_0^\Theta$ and $\eta_{\theta_0^i,0}^X = N(m_{\theta_0^i,0}, \Sigma_{\theta_0^i,0})$;

end

for $p \leftarrow 0$ **to** $n - 1$ **do**

/* Selection of Kalman filters */

Sample $I_p = (I_p^i)_{i=1}^{N_1}$ according to a multinomial distribution with probability proportional to $\left(\bar{G}_p(\theta_p^k, \eta_{\theta_0^k,p}^X)\right)_{k=1}^{N_1}$ given by (14) ;

for $i \leftarrow 1$ **to** N_1 **do**

/* Updating step for each Kalman filter */

$$\begin{cases} \hat{m}_{\theta_p^i, p} = m_{\theta_p^i, p} + K_{\theta_p^i, p}(Y_n - C_{\theta_p^i, p} m_{\theta_p^i, p}) \\ \hat{\Sigma}_{\theta_p^i, p} = (I - K_{\theta_p^i, p} C_{\theta_p^i, p}) \Sigma_{\theta_p^i, p} \end{cases};$$

/* Mutation of each island */

Sample independently θ_{p+1}^i according to $M_{p+1}^\Theta(\theta_{p+1}^i, .)$;

/* Prediction step for each Kalman filter */

$$\begin{cases} m_{\theta_{p+1}^i, p+1} = A_{\theta_{p+1}^i, p+1} \hat{m}_{\theta_p^i, p} + a_{\theta_{p+1}^i, p+1} \\ \Sigma_{\theta_{p+1}^i, p+1} = A_{\theta_{p+1}^i, p+1} \hat{\Sigma}_{\theta_p^i, p} (A_{\theta_{p+1}^i, p+1})^T + B_{\theta_{p+1}^i, p+1} \Sigma_{p+1}^X (B_{\theta_{p+1}^i, p+1})^T \end{cases}$$

end

$p \leftarrow p + 1$

end

Algorithm 1: Interacting Kalman Filter - IKF

Secondly, when the non linear Equation 6 can be solved analytically i.e. when one has access to the exact measure $\eta_{\theta_{0:n},n}^X$, one can apply a simple interacting particle model as described in Figure 1, where each particle corresponds to the pair: parameter and exact measure.

However, in most cases, this equation cannot be solved analytically, so that an additional approximation is needed in order to estimate the measure $\eta_{\theta_{0:n},n}^X$ for each $i \in [1, N_1]$. The next subsection is dedicated to the derivation of an algorithm to deal with this constraint.

2.2 Labeled island particle model

To tackle the case where $\eta_{\theta_{0:n},n}^X$, $i \in [1, N_1]$ is not analytically known, the idea consists in using a particle estimation of $\eta_{\theta_{0:n},n}^X$ inside the previous interacting particle model. The ensuing algorithm will be called *labeled island particle model* in reference to the island particle model developed in [7], even if in the present case, each island i have a label θ_n^i whose evolution is given by the Markov kernel M_n^Θ . The labeled island particle model consists in associating to each term of the sequence $(\theta_n^i)_{i=1}^{N_1}$ a sub N_2 -interacting particle system. We call sub N_2 -interacting particle system associated with the sequence $((G_{\theta_n^i,n}, M_{\theta_n^i,n}^X))_{n \in \mathbb{N}}$ and the initial distribution $\eta_{\theta_0^i,0}^X$, the sequence of non-homogeneous Markov chain $(\xi_n^{i,j})_{j=1}^{N_2}$ taking value in the product space E_n^{X,N_2} , that is :

$$\xi_n^{i,[N_2]} \triangleq (\xi_n^{i,j})_{j=1}^{N_2} \triangleq (\xi_n^{i,1}, \dots, \xi_n^{i,N_2}) \in E_n^{X,N_2} \triangleq \underbrace{E_n^X \times \dots \times E_n^X}_{N_2 \text{ times}}.$$

The initial state of the Markov chain $(\xi_0^{i,j})_{j=1}^{N_2}$ consists in sampling N_2 independent random variables with common distribution $\eta_{\theta_0^i,0}^X$.

The interacting particle system, denoted by $(\xi_n^{i,j})_{j=1}^{N_2}$, explore the state space \mathbb{E}_n^X and with the dynamic given to it, empirically sample the law $\eta_{\theta_{0:n},n}^X$.

Denoting the empirical measure

$$\eta_{\theta_{0:n},n}^{X,N_2} \triangleq \frac{1}{N_2} \sum_{j=1}^{N_2} \delta_{\xi_n^{i,j}}, \quad (15)$$

the elementary transition of the process $\xi_n^{i,[N_2]}$ from \mathbb{E}_n^{X,N_2} to \mathbb{E}_{n+1}^{X,N_2} is given for any $x_n^{[N_2]} = (x_n^1, \dots, x_n^{N_2}) \in \mathbb{E}_n^{X,N_2}$ by

$$\begin{aligned} \mathbb{P}_{\eta_{\theta_0}^X}^{N_2} \left(\xi_{n+1}^{i,[N_2]} \in dx_{n+1}^{[N_2]} \mid \xi_n^{i,[N_2]} \right) &\triangleq \prod_{j=1}^{N_2} \Phi_{n+1}^X \left((\theta_n^i, \theta_{n+1}^i), \eta_{\theta_{0:n},n}^{X,N_2} \right) (dx_{n+1}^j) \\ &= \prod_{j=1}^{N_2} \Psi_{\theta_n^i,n}^X(\eta_{\theta_{0:n},n}^{X,N_2}) M_{\theta_{n+1}^i,n+1}^X(\xi_n^{i,j}, dx_{n+1}^j) \quad \text{using (6)} \\ &= \prod_{j=1}^{N_2} \sum_{k=1}^{N_2} \frac{G_{\theta_n^i,n}(\xi_n^{i,k})}{\sum_{\ell=1}^{N_2} G_{\theta_n^i,n}(\xi_n^{i,\ell})} M_{\theta_{n+1}^i,n+1}^X(\xi_n^{i,k}, dx_{n+1}^j) \quad \text{by (5).} \end{aligned}$$

Define the mapping $\tilde{\Phi}_n^X$ by

$$\begin{aligned} \tilde{\Phi}_n^X : \mathbb{E}_{n-1}^\Theta \times \mathbb{E}_n^\Theta \times \mathcal{P}(\mathbb{E}_{n-1}^X) &\rightarrow \mathcal{P}(\mathbb{E}_n^X) \\ ((u, v), \nu) &\mapsto \prod_{j=1}^{N_2} \Phi_n^X((u, v), \nu)(dx_n^j), \end{aligned}$$

then

$$\eta_{\theta_{0:n+1},n+1}^{X,N_2} = \tilde{\Phi}_{n+1}^X \left((\theta_n^i, \theta_{n+1}^i), \eta_{\theta_{0:n},n}^{X,N_2} \right). \quad (16)$$

So, the evolution of the particle swarm $\xi_n^{i,[N_2]}$ consists in two steps: a selection and a mutation. In the selection step, the particles are selected multinomially with probability proportional to their potentials $(G_{\theta_n^i,n}(\xi_n^{i,j}))_{j=1}^{N_2}$. Then the mutation step is performed independently using the kernel $M_{\theta_{n+1}^i,n+1}^X$. Hence, at each iteration $n \in \mathbb{N}$, the empirical measure $\eta_{\theta_{0:n},n}^{X,N_2}$ approximates $\eta_{\theta_{0:n},n}^X$ when N_2 tends to ∞ . Replacing $\eta_{\theta_{0:n},n}^X$ by $\eta_{\theta_{0:n},n}^{X,N_2}$ inside the first algorithm presented, one gets a nested particle model named *labeled island particle model*.

In order to derive precisely this algorithm, first introduce the following sequence \tilde{X}_n on $\bar{\mathbb{E}}_n = \mathbb{E}_n^\Theta \times \mathcal{P}(\mathbb{E}_n^X)$, defined by $\tilde{X}_n \triangleq (\Theta_n, \eta_{\theta_{0:n},n}^{X,N_2})$, i.e. the couple environment and empirical measure of the process X_n conditionally on $\Theta_{0:n}$, where $\eta_{\theta_{0:n},n}^{X,N_2} \triangleq \sum_{j=1}^{N_2} \delta_{\xi_n^j}/N_2$.

Proposition 2.1. \tilde{X}_n is a $\bar{\mathbb{E}}_n$ -Markov chain with transition kernel \tilde{M}_n defined for every function $\bar{f}_n \in \mathcal{B}_b(\bar{\mathbb{E}}_n)$ and $(u, \nu) \in \bar{\mathbb{E}}_n$ by

$$\tilde{M}_n(\bar{f}_n)(u, \nu) = \int_{\mathbb{E}_n^\Theta} M_n^\Theta(u, dv) \bar{f}_n(v, \tilde{\Phi}_n^X((u, v), \nu)), \quad (17)$$

where $\tilde{\Phi}_n^X$ is defined in (16), and with initial distribution $\tilde{\eta}_0 \in \mathcal{P}(\bar{\mathbb{E}}_0)$ given by

$$\tilde{\eta}_0(d(u, \nu)) \triangleq \eta_0^\Theta(du) \delta_{\eta_{\theta_0,0}^{X,N_2}}(d\nu).$$

Proof. Let $\sigma(\tilde{X}_0, \dots, \tilde{X}_n)$ stands for the σ -algebra generated by the random variables \tilde{X}_p , $0 \leq p \leq n$. For all $\bar{f}_n \in \mathcal{B}_b(\bar{\mathbb{E}}_n)$:

$$\begin{aligned} \mathbb{E}_{\tilde{\eta}_0}[\bar{f}_n(\tilde{X}_n) | \sigma(\tilde{X}_0, \dots, \tilde{X}_{n-1})] \\ = \mathbb{E}_{\tilde{\eta}_0}[\bar{f}_n(\Theta_n, \eta_{\Theta_{0:n}, n}^{X, N_2}) | \sigma(\tilde{X}_0, \dots, \tilde{X}_{n-1})] \\ = \mathbb{E}_{\tilde{\eta}_0}[\bar{f}_n(\Theta_n, \tilde{\Phi}_n^X((\Theta_{n-1}, \Theta_n), \eta_{\Theta_{0:n-1}, n-1}^{X, N_2}) | \sigma(\tilde{X}_0, \dots, \tilde{X}_{n-1})] \quad \text{by (16).} \end{aligned}$$

Recalling that $\tilde{X}_{n-1} = (\Theta_{n-1}, \eta_{\Theta_{0:n-1}, n-1}^{X, N_2})$, one can conclude that

$$\begin{aligned} \mathbb{E}_{\tilde{\eta}_0}[\bar{f}_n(\tilde{X}_n) | \sigma(\tilde{X}_0, \dots, \tilde{X}_{n-1})] \\ = \mathbb{E}_{\tilde{\eta}_0}[\bar{f}_n(\tilde{X}_n) | \tilde{X}_{n-1}] \\ = \int_{\mathbb{E}_n^\Theta} \bar{f}_n(\theta_n, \tilde{\Phi}_n^X((\Theta_{n-1}, \theta_n), \eta_{\Theta_{0:n-1}, n-1}^{X, N_2}) M_n^\Theta(\Theta_{n-1}, d\theta_n). \end{aligned}$$

□

To the Markov chain \tilde{X}_n , one may associate the Feynman-Kac distribution defined for every $\bar{f}_n \in \mathcal{B}_b(\bar{\mathbb{E}}_n)$ by

$$\tilde{\eta}_n(\bar{f}_n) \triangleq \tilde{\gamma}_n(\bar{f}_n)/\tilde{\gamma}_n(\mathbb{1}), \quad (18)$$

where $\tilde{\gamma}_n$ is defined such that

$$\tilde{\gamma}_n(\bar{f}_n) \triangleq \mathbb{E}_{\tilde{\eta}_0} \left[\bar{f}_n(\tilde{X}_n) \prod_{p=0}^{n-1} \bar{G}_p(\tilde{X}_p) \right],$$

with \bar{G}_p defined in (10).

In a similar way to $\bar{\eta}_n$, the measure $\tilde{\eta}_n$ satisfies a recursive equation $\tilde{\eta}_n = \bar{\Psi}_{n-1}(\tilde{\eta}_{n-1})\tilde{M}_n$, with $\bar{\Psi}_{n-1}$ the application defined in Proposition 1.3. This non linear equation can be rewritten as

$$\tilde{\eta}_n = \tilde{\Phi}_n(\tilde{\eta}_{n-1}), \quad (19)$$

where the mapping $\tilde{\Phi}_n$ is defined as follows :

$$\begin{aligned} \tilde{\Phi}_n : \mathcal{P}(\bar{\mathbb{E}}_{n-1}) &\rightarrow \mathcal{P}(\bar{\mathbb{E}}_n) \\ \eta &\mapsto \bar{\Psi}_{n-1}(\eta)\tilde{M}_n. \end{aligned} \quad (20)$$

As in Section 2.1, when this equation cannot be solved analytically one may use a particle model to approximate the probability measure $\tilde{\eta}_n$. In this case, the particles $\{\tilde{X}_n^i \triangleq (\theta_n^i, \eta_{\theta_{0:n}^i, n}^{X, N_2}), i \in \llbracket 1, N_1 \rrbracket\}$, would be testing points on the state space $\bar{\mathbb{E}}_n$, for $(N_1, N_2) \in (\mathbb{N}^*)^2$. These particles explore the state space $\bar{\mathbb{E}}_n$ and their dynamics empirically sample the law $\tilde{\eta}_n$ when N_1 gets large. An interacting particle system associated with the couple (\bar{G}_n, \tilde{M}_n) and the initial distribution $\tilde{\eta}_0$, is a sequence of non-homogeneous Markov chain, $\tilde{X}_n^{[N_1]}$, taking value in the product space $\bar{\mathbb{E}}_n^{N_1}$, defined by

$$\tilde{X}_n^{[N_1]} \triangleq (\tilde{X}_n^i)_{i=1}^{N_1} = (\tilde{X}_n^1, \dots, \tilde{X}_n^{N_1}) \in \bar{\mathbb{E}}_n^{N_1}.$$

The initial state of the Markov chain $\tilde{X}_0^{[N_1]}$ consists in N_1 independent random variables with common distribution $\tilde{\eta}_0$. Denote by $\tilde{\eta}_n^{N_1}$ the empirical measure at time n , which is defined by

$$\tilde{\eta}_n^{N_1} \triangleq \frac{1}{N_1} \sum_{i=1}^{N_1} \delta_{\tilde{X}_n^i}. \quad (21)$$

The elementary probability transition, is given for any $\bar{x}_{n+1}^{[N_1]} \in \bar{\mathcal{E}}_{n+1}^{N_1}$ by

$$\mathbb{P}_{\tilde{\eta}_0}^{N_1}(\tilde{X}_{n+1}^{[N_1]} \in d\bar{x}_{n+1}^{[N_1]} | \tilde{X}_n^{[N_1]}) = \prod_{i=1}^{N_1} \bar{\Psi}_n(\tilde{\eta}_n^{N_1}) \tilde{M}_{n+1}(\tilde{X}_n^i, d\bar{x}_{n+1}^i).$$

The particle evolution is summarized in Figure 2 where by definitions (10) and (15),

$$\bar{G}_n(\tilde{X}_n^i) = \frac{1}{N_2} \sum_{j=1}^{N_2} G_n(\theta_n^i, \xi_n^{i,j}) = \frac{1}{N_2} \sum_{j=1}^{N_2} G_{\theta_n^i, n}(\xi_n^{i,j}).$$

The ensuing algorithm is described in Algorithm 2.

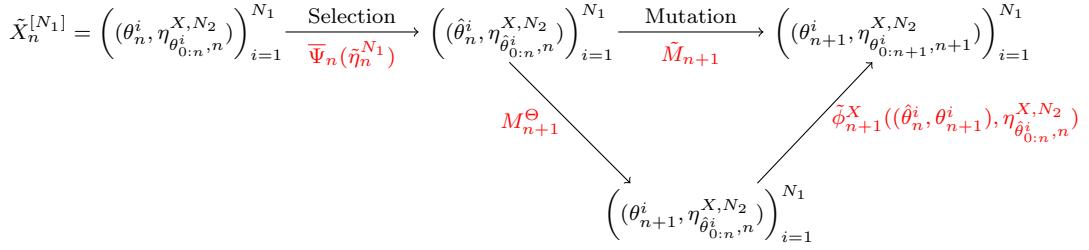


Figure 2: Evolution scheme of the labeled island particle model.

Data: $\tilde{\eta}_0$, $(\tilde{M}_p)_{p=0}^n$ and $(\bar{\Psi}_p)_{p=0}^n$

Result: Particle approximation of $\tilde{\eta}_n$

```

/* Initialization */
```

for $i \leftarrow 1$ **to** N_1 **do**

Sample $\tilde{X}_0^i = (\theta_0^i, \eta_{\theta_0^i, 0}^{X, N_2}) \sim \tilde{\eta}_0$, that is

$$\theta_0^i \stackrel{i.i.d.}{\sim} \eta_0^\Theta, \xi_0^{i,j} \stackrel{i.i.d.}{\sim} \eta_{\theta_0^i, 0}^X \text{ and } \eta_{\theta_0^i, 0}^{X, N_2} = \frac{1}{N_2} \sum_{j=1}^{N_2} \delta_{\xi_0^{i,j}}.$$

end

for $p \leftarrow 0$ **to** $n - 1$ **do**

/* Selection of islands */

Sample $I_p = (I_p^i)_{i=1}^{N_1}$ according to a multinomial distribution with probability proportional to $\left(\frac{1}{N_2} \sum_{j=1}^{N_2} G_p(\theta_p^i, \xi_p^{i,j}) \right)_{i=1}^{N_1}$;

/* Selection of particles inside each island */

Sample $J_p^i = (J_p^{i,j})_{j=1}^{N_2}$ according to a multinomial distribution with probability proportional to $\left(G_p(\theta_p^{I_p^i}, \xi_p^{I_p^i, j}) \right)_{j=1}^{N_2}$;

/* Mutation of each island */

Sample independently θ_{p+1}^i according to $M_{p+1}^\Theta(\theta_p^{I_p^i}, \cdot)$;

/* Mutation of particles */

Sample $\xi_{p+1}^{i,j}$ according to $M_{\theta_{p+1}^i, p+1}^X(\xi_p^{I_p^i, J_p^{i,j}}, \cdot)$;

end

end

$p \leftarrow p + 1$

end

Algorithm 2: Labeled island particle algorithm

For every $n \geq 0$, $\tilde{\eta}_n^{N_1}$ is an estimator of $\tilde{\eta}_n$, obtained through the labeled island particle model, i.e. for every $\bar{f}_n \in \mathcal{B}_b(\bar{\mathbb{E}}_n)$,

$$\tilde{\eta}_n^{N_1}(\bar{f}_n) = \frac{1}{N_1} \sum_{i=1}^{N_1} \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^{X, N_2})$$

converges to $\tilde{\eta}_n(\bar{f}_n)$ when $N_1 \rightarrow +\infty$.

3 \mathbb{L}^p bounds

We are interested in this section in the \mathbb{L}^p bounds of the difference between the estimator $\tilde{\eta}_n^{N_1}$ and the measure $\bar{\eta}_n$. To get these bounds we will use several notations. We define them before going further.

3.1 Notations

Let $(\mathsf{E}, \mathcal{E})$ be a measurable space. For a real-valued measurable function $h \in \mathcal{B}_b(\mathsf{E})$, we denote the oscillator norm $\text{osc}(h) \triangleq \sup_{(x,x') \in \mathsf{E}^2} |h(x) - h(x')|$, and $\text{Osc}_1(\mathsf{E})$ the convex set of \mathcal{E} -measurable functions with oscillations less than one. The sup norm of h is noted $\|h\|_\infty \triangleq \sup_{x \in \mathsf{E}} |h(x)|$ and the \mathbb{L}^p -norm $\|\cdot\|_p$. $\mathcal{B}_1(\mathsf{E}) \subset \mathcal{B}_b(\mathsf{E})$ refers to the set of functions whose sup norm is less than one. For two probability measures $(\mu, \eta) \in \mathcal{P}(\mathsf{E})^2$, the Zolotarev semi-norm $\|\cdot\|_{\mathfrak{F}}$ attached to \mathfrak{F} a countable collection of bounded measurable functions in $\mathcal{B}_1(\mathsf{E})$ is defined by

$$\|\mu - \eta\|_{\mathfrak{F}} \triangleq \sup_{f \in \mathfrak{F}} |\mu(f) - \eta(f)|.$$

To measure the size of a given class \mathfrak{F} , one considers the covering numbers $\mathcal{N}(\varepsilon, \mathfrak{F}, \mathbb{L}^p(\mu))$ defined as the minimal number of $\mathbb{L}^p(\mu)$ -balls of radius $\varepsilon > 0$ needed to cover \mathfrak{F} . Let $\mathcal{N}(\varepsilon, \mathfrak{F})$ and $I(\mathfrak{F})$ denote respectively the uniform covering numbers and entropy integral given by

$$\mathcal{N}(\varepsilon, \mathfrak{F}) \triangleq \sup_{\mu \in \mathcal{P}(\mathsf{E})} \mathcal{N}(\varepsilon, \mathfrak{F}, \mathbb{L}^2(\mu)) \quad (22)$$

$$I(\mathfrak{F}) \triangleq \int_0^1 \sqrt{\log(1 + \mathcal{N}(\varepsilon, \mathfrak{F}))} d\varepsilon. \quad (23)$$

Let \wedge denote the minimum operator and \vee denote the maximum operator. For a kernel M defined on E , the Dobrushin coefficient of M is

$$\beta(M) \triangleq \sup_{f \in \text{Osc}_1(\mathsf{E})} \text{osc}(M(f)).$$

Let $(d(n))_{n \geq 0}$ be a sequence defined for every $m \geq 0$ by

$$\begin{cases} d(2m)^{2m} \triangleq (2m)_m 2^{-m} \\ d(2m+1)^{2m+1} \triangleq \frac{(2m+1)_m 2^{m+1}}{\sqrt{m+1/2}} 2^{-m+1/2}, \end{cases}$$

where for any positive integers $(p, q) \in (\mathbb{N}^*)^2$, $(q+p)_p \triangleq (q+p)!/q!$.

For $n \in \mathbb{N}$, introduce the Feynman-Kac semi-groups \overline{Q}_n (resp. $Q_{\theta_{n-1:n}, n}^X$) such that for all $(\bar{x}_n, \bar{x}_{n+1}) \in \bar{\mathsf{E}}_n \times \bar{\mathsf{E}}_{n+1}$ (resp. $(x_n, x_{n+1}) \in \mathsf{E}_n^X \times \mathsf{E}_{n+1}^X$),

$$\overline{Q}_{n+1}(\bar{x}_n, d\bar{x}_{n+1}) \triangleq \overline{G}_n(\bar{x}_n) \overline{M}_{n+1}(\bar{x}_n, d\bar{x}_{n+1}),$$

$$\left(\text{resp. } Q_{\theta_{n:n+1}, n+1}^X(x_n, dx_{n+1}) \triangleq G_{\theta_n, n}(x_n) M_{\theta_{n+1}, n+1}^X(x_n, dx_{n+1}) \right).$$

For every $(p, n) \in (\mathbb{N})^2$ such that $p < n$, set

$$\overline{Q}_{p,n} \triangleq \overline{Q}_{p+1} \dots \overline{Q}_n, \quad \text{and} \quad \overline{P}_{p,n} \triangleq \overline{Q}_{p,n}/\overline{Q}_{p,n}(\mathbb{1}),$$

$$\left(\text{resp. } Q_{\theta_{p:n}, p, n}^X \triangleq Q_{\theta_{p:p+1}, p+1}^X \dots Q_{\theta_{n-1:n}, n}^X \text{ and } P_{\theta_{p:n}, p, n}^X(f_n) \triangleq Q_{\theta_{p:n}, p, n}^X(f_n)/Q_{\theta_{p:n}, p, n}^X(\mathbb{1}) \right),$$

and set the normalizing constant

$$\overline{G}_{p,n} \triangleq \overline{Q}_{p,n}(\mathbb{1}), \quad (\text{resp. } G_{\theta_{p:n}, p, n} \triangleq Q_{\theta_{p:n}, p, n}^X(\mathbb{1})).$$

Finally, set

$$\bar{g}_{p,n} \triangleq \sup_{(\bar{x}_p, \bar{y}_p) \in (\bar{\mathsf{E}}_p)^2} \frac{\overline{G}_{p,n}(\bar{x}_p)}{\overline{G}_{p,n}(\bar{y}_p)}, \quad \left(\text{resp. } g_{\theta_{p:n}, p, n} \triangleq \sup_{(x_p, y_p) \in (\mathsf{E}_p^X)^2} \frac{G_{\theta_{p:n}, p, n}(x_p)}{G_{\theta_{p:n}, p, n}(y_p)} \right).$$

In order to study the difference between $\tilde{\eta}_n^{N_1}$ and $\bar{\eta}_n$, we use several results taken from [4]. Then, we will always assume that for all $n \in \mathbb{N}$, the potential functions $G_{\theta_n, n}$ defined in (3) satisfy the following condition (G_θ):

there exists a sequence of strictly positive number $\epsilon_n(G_\theta) \in (0, 1]$ such that for any $(x_n, y_n) \in (\mathbb{E}_n^X)^2$:

$$G_{\theta_n, n}(x_n) \geq \epsilon_n(G_\theta) G_{\theta_n, n}(y_n) > 0 \quad (G_\theta)$$

Therefore, for all $n \in \mathbb{N}$, the potential functions \bar{G}_n satisfy the following condition (\bar{G}):
there exists a sequence of strictly positive number $\epsilon_n(\bar{G}) \in (0, 1]$ such that for any $(\bar{x}_n, \bar{y}_n) \in (\bar{\mathbb{E}}_n)^2$:

$$\bar{G}_n(\bar{x}_n) \geq \epsilon_n(\bar{G}) \bar{G}_n(\bar{y}_n) > 0 \quad (\bar{G})$$

Moreover we always assume that the collection of distributions $(\bar{M}_{n+1}(\bar{x}_n, .))_{\bar{x}_n \in \bar{\mathbb{E}}_n}$ are absolutely continuous with one another. That is for every $n \geq 0$ and $(\bar{x}_n, \bar{y}_n) \in (\bar{\mathbb{E}}_n)^2$, one has

$$\bar{M}_{n+1}(\bar{x}_n, .) \ll \bar{M}_{n+1}(\bar{y}_n, .).$$

In addition, we assume that the collection of distributions $(M_{\theta_{n+1}, n+1}^X(x_n, .))_{x_n \in \mathbb{E}_n^X}$ are absolutely continuous with one another. That is for every $n \geq 0$, $\theta_{n+1} \in \mathbb{E}_{n+1}^\Theta$ and $(x_n, y_n) \in (\mathbb{E}_n^X)^2$, one has :

$$M_{\theta_{n+1}, n+1}^X(x_n, .) \ll M_{\theta_{n+1}, n+1}^X(y_n, .).$$

Note that for time homogeneous models on finite spaces condition those conditions are met as soon as the Markov chain is aperiodic and irreducible. Some examples are illustrated by typical examples in [4].

3.2 \mathbb{L}^p bound

Consider that for all $n \in \mathbb{N}$, the product space $\bar{\mathbb{E}}_n = \mathbb{E}_n^\Theta \times \mathcal{P}(\mathbb{E}_n^X)$ is equipped with the norm $\|\cdot\|_{\bar{\mathbb{E}}_n}$ such that for all $(u, v) \in (\mathbb{E}_n^\Theta)^2$ and $(\nu, \eta) \in (\mathcal{P}(\mathbb{E}_n^X))^2$,

$$\|(u, \eta) - (v, \nu)\|_{\bar{\mathbb{E}}_n} = |u - v| + \|\eta - \nu\|_{\mathfrak{F}_n}.$$

where \mathfrak{F}_n is a countable collection of functions in $\mathcal{B}_1(\mathbb{E}_n^X)$.

Theorem 3.1. *For any $p \in \mathbb{N}^*$, $n \in \mathbb{N}$, let $\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)$ be a k_n -Lipschitz function. Assume that for any $\theta_n \in \mathbb{E}_n^\Theta$, the kernel transition $M_{\theta_n, n}^X$ can be written as $M_{\theta_n, n}^X(x_{n-1}, dx_n) = m_{\theta_n, n}^X(x_{n-1}, x_n)p_{\theta_n, n}(dx_n)$ for some measurable function $m_{\theta_n, n}^X$ on $\mathbb{E}_{n-1}^X \times \mathbb{E}_n^X$ and some probability measure $p_{\theta_n, n} \in \mathcal{P}(\mathbb{E}_n^X)$. Furthermore, assume that there exists a collection of mappings $\alpha_{\theta_n, n}$ on \mathbb{E}_n^X such that*

$$\sup_{x_{n-1} \in \mathbb{E}_{n-1}^X} |\log m_{\theta_n, n}^X(x_{n-1}, x_n)| \leq \alpha_{\theta_n, n}(x_n)$$

with $p_{\theta_n, n}(e^{3\alpha_{\theta_n, n}}) < \infty$.

Then, the \mathbb{L}^p error is bounded by

$$\|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p \leq k_n \frac{a(p)}{\sqrt{N_2}} (I(\mathfrak{F}_n) + b(n)) + 2 \frac{d(p)}{\sqrt{N_1}} \sum_{q=0}^n \bar{g}_{q, n} \beta(\bar{P}_{q, n}), \quad (24)$$

where the sequence $d(n)$ is defined in (24), $I(\mathfrak{F}_n)$ is defined in (23), $(b(n))_{n \geq 0}$ is defined by

$$b(0) = 0 \text{ and } b(n+1) \leq g_{\theta_n, n} p_{\theta_{n+1}, n+1} (e^{3\alpha_{\theta_{n+1}, n+1}}) \sum_{q=0}^n g_{\theta_{q:n}, q, n} \beta(P_{\theta_{q:n}, q, n}),$$

and $a(n)$ is a sequence such that for all $n \in \mathbb{N}^*$, $a(n) \leq c[n/2]!$ with c a universal constant.

Proof. Let $\bar{f}_n \in \text{Osc}_1(\bar{\mathbf{E}}_n)$ be a k_n -Lipschitz function, and apply triangular inequality:

$$\|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p \leq \|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n^{N_1}(\bar{f}_n)\|_p + \|\bar{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p,$$

where $\bar{\eta}_n$ is defined in (9). Then using Theorem 7.4.4 from [4], one can bound the second term

$$\|\bar{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p \leq 2 \frac{d(p)}{\sqrt{N_1}} \sum_{q=0}^n \bar{g}_{q,n} \beta(\bar{P}_{q,n}).$$

Therefore, in order to bound the first term, use the definitions of $\tilde{\eta}_n^{N_1}$ and $\bar{\eta}_n^{N_1}$ in (21) and (13) respectively :

$$\begin{aligned} \|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n^{N_1}(\bar{f}_n)\|_p &= \mathbb{E}_{\bar{\eta}_0} \left[|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n^{N_1}(\bar{f}_n)|^p \right]^{1/p} \\ &= \mathbb{E}_{\bar{\eta}_0} \left[\left| \frac{1}{N_1} \sum_{i=1}^{N_1} \left\{ \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^{X, N_2}) - \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^X) \right\} \right|^p \right]^{1/p} \\ &\leq \mathbb{E}_{\bar{\eta}_0} \left[\left(\frac{1}{N_1} \sum_{i=1}^{N_1} |\bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^{X, N_2}) - \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^X)| \right)^p \right]^{1/p} \end{aligned}$$

As \bar{f}_n is k_n -Lipschitz, it follows

$$\begin{aligned} \left| \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^{X, N_2}) - \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^X) \right| &\leq k_n \left\| (\theta_n^i, \eta_{\theta_{0:n}^i, n}^{X, N_2}) - (\theta_n^i, \eta_{\theta_{0:n}^i, n}^X) \right\|_{\bar{\mathbf{E}}_n} \\ &\leq k_n \left\| \eta_{\theta_{0:n}^i, n}^{X, N_2} - \eta_{\theta_{0:n}^i, n}^X \right\|_{\mathfrak{F}_n}. \end{aligned}$$

where \mathfrak{F}_n is a countable collection of functions in $\mathcal{B}_1(\mathbf{E}_n^X)$ Therefore one gets

$$\|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n^{N_1}(\bar{f}_n)\|_p \leq \mathbb{E}_{\bar{\eta}_0} \left[\left(\frac{1}{N_1} \sum_{i=1}^{N_1} k_n \left\| \eta_{\theta_{0:n}^i, n}^{X, N_2} - \eta_{\theta_{0:n}^i, n}^X \right\|_{\mathfrak{F}_n} \right)^p \right]^{1/p}.$$

Denoting by θ^* the value at which the maximum of the Zolotarev semi-norms for $i \in \llbracket 1, N_1 \rrbracket$ is reached, yields to

$$\|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n^{N_1}(\bar{f}_n)\|_p \leq k_n \mathbb{E}_{\bar{\eta}_0} \left[\left\| \eta_{\theta_{0:n}^*, n}^{X, N_2} - \eta_{\theta_{0:n}^*, n}^X \right\|_{\mathfrak{F}_n}^p \right]^{1/p}.$$

But, according to [[4], Corollary 7.4.4] (p. 247),

$$\mathbb{E}_{\bar{\eta}_0} \left[\left\| \eta_{\theta_{0:n}^*, n}^{X, N_2} - \eta_{\theta_{0:n}^*, n}^X \right\|_{\mathfrak{F}_n}^p \right]^{1/p} \leq \frac{a(p)}{\sqrt{N_2}} (I(\mathfrak{F}_n) + b(n)),$$

where $I(\mathfrak{F}_n)$ is the entropy of \mathfrak{F}_n defined in (23). Finally one ends up with

$$\|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n^{N_1}(\bar{f}_n)\|_p \leq k_n \frac{a(p)}{\sqrt{N_2}} (I(\mathfrak{F}_n) + b(n)) + 2 \frac{d(p)}{\sqrt{N_1}} \sum_{q=0}^n \bar{g}_{q,n} \beta(\bar{P}_{q,n}).$$

□

3.3 Time uniform bound

Before stating the uniform estimate we define the following two additional conditions.

There exists some integer $\bar{m} \geq 1$ and some numbers $\varepsilon_n(\bar{M}) \in]0, 1[$ such that for $n \in \mathbb{N}$ and $(\bar{x}_n, \bar{y}_n) \in \bar{\mathbb{E}}_n^2$, one has :

$$\bar{M}_{n,n+\bar{m}}(\bar{x}_n, .) \triangleq \bar{M}_{n+1} \dots \bar{M}_{n+\bar{m}}(\bar{x}_n, .) \geq \varepsilon_n(\bar{M}) \bar{M}_{n,n+\bar{m}}(\bar{y}_n, .) \quad ((\bar{M})_{\bar{m}})$$

For all $i \in \llbracket 1, N_1 \rrbracket$, there exists some integer $m_i \geq 1$ and some numbers $\varepsilon_n(M_{\theta^i}^X) \in]0, 1[$ such that for $n \in \mathbb{N}$, $(x_n, y_n) \in (\mathbb{E}_n^X)^2$, and $\theta_{n+1:n+m_i}^i \in \mathbb{E}_{n+1}^\Theta \times \dots \times \mathbb{E}_{n+m}^\Theta$, one has :

$$\begin{aligned} M_{\theta_{n+1:n+m_i}^i, n, n+m_i}^X(x_n, .) &\triangleq M_{\theta_{n+1}^i, n+1}^X \dots M_{\theta_{n+m_i}^i, n+m_i}^X(x_n, .) \\ &\geq \varepsilon_n(M_{\theta^i}^X) M_{\theta_{n+1:n+m_i}^i, n, n+m}^X(y_n, .) \quad ((M_{\theta^i}^X)_{m_i}) \end{aligned}$$

Theorem 3.2. Suppose that conditions (\bar{G}) , $((\bar{M})_{\bar{m}})$ are met for some integer $\bar{m} \geq 1$ and some pair parameters $(\varepsilon_n(\bar{G}), \varepsilon_n(\bar{M}))$ and set $\epsilon(\bar{G}) \triangleq \wedge_{n \geq 0} \varepsilon_n(\bar{G})$ and $\varepsilon(\bar{M}) \triangleq \wedge_{n \geq 0} \varepsilon_n(\bar{M})$.

Moreover, assume that for all $i \in \llbracket 1, N_1 \rrbracket$ conditions (G_θ) and $((M_{\theta^i}^X)_{m_i})$ hold true for some sequence of integer m_i and some pair parameters $(\varepsilon_n(G_{\theta^i}), \varepsilon_n(M_{\theta^i}^X))$ and set $\epsilon(G_{\theta^i}) \triangleq \wedge_{n \geq 0} \varepsilon_n(G_{\theta^i})$ and $\varepsilon(M_{\theta^i}^X) \triangleq \wedge_{n \geq 0} \varepsilon_n(M_{\theta^i}^X)$. Set $m \triangleq \vee_i m_i$.

Further assume that for all $n \geq 0$ and $\theta_n^i \in \mathbb{E}_n^\Theta$ the kernel transition $M_{\theta_n^i, n}^X$ has the form $M_{\theta_n^i, n}^X(x_{n-1}, dx_n) = m_{\theta_n^i, n}^X(x_{n-1}, x_n) p_{\theta_n^i, n}(dx_n)$ for some measurable function $m_{\theta_n^i, n}^X$ on $\mathbb{E}_{n-1}^X \times \mathbb{E}_n^X$ and some probability measure $p_{\theta_n^i, n} \in \mathcal{P}(\mathbb{E}_n^X)$.

Also assume that $\sup_{x_{n-1} \in \mathbb{E}_{n-1}^X} |\log m_{\theta_n^i, n}^X(x_{n-1}, x_n)| \leq \alpha_{\theta_n^i, n}(x_n)$ with $p_{\theta_n^i, n}(e^{3\alpha_{\theta_n^i, n}}) < \infty$ for some collection of mappings $\alpha_{\theta_n^i, n}$ on \mathbb{E}_n^X , and set :

$$p_{\theta^i}(e^{3\alpha_{\theta^i}}) \triangleq \sup_{n \geq 0} p_{\theta_n^i, n}(e^{3\alpha_{\theta_n^i, n}}) < \infty \quad \text{and} \quad p_\theta(e^{3\alpha_\theta}) \triangleq \vee_i p_{\theta^i}(e^{3\alpha_{\theta^i}}).$$

Then for any $p \in \mathbb{N}^*$, any k_n -Lipschitz functions $\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)$ one has :

$$\sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} \|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p \leq \frac{2d(p)\bar{m}}{\sqrt{N_1}\varepsilon(\bar{M})^3\epsilon(\bar{G})^{2\bar{m}-1}} + \frac{k a(p)}{\sqrt{N_2}} \left(I + \frac{mp_\theta(e^{3\alpha_\theta})}{\varepsilon(M_\theta^X)^3\epsilon(G_\theta)^{2m}} \right)$$

with

$$k = \sup_n k_n, \quad \varepsilon(M_\theta^X) = \wedge_i \varepsilon(M_{\theta^i}^X) > 0, \quad \epsilon(G_\theta) = \wedge_i \epsilon(G_{\theta^i}) > 0, \quad I \triangleq \sup_{n \geq 0} I(\mathfrak{F}_n) < \infty.$$

Proof.

$$\begin{aligned} &\sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} \|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p \\ &\leq \sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} (\|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n^{N_1}(\bar{f}_n)\|_p + \|\bar{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p) \\ &\leq \sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} \|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n^{N_1}(\bar{f}_n)\|_p + \sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} \|\bar{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p \end{aligned}$$

From [4], Theorem 7.4.4] (p. 247), one has

$$\sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} \|\bar{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p \leq \frac{2d(p)\bar{m}}{\sqrt{N_1}\varepsilon(\bar{M})^3\epsilon(\bar{G})^{2\bar{m}-1}},$$

since conditions (\bar{G}) and $((\bar{M})_{\bar{m}})$ hold true. Then it follows that the only term one has to work on is the following.

$$\begin{aligned} & \sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} \|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n^{N_1}(\bar{f}_n)\|_p \\ &= \sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} \left\| \frac{1}{N_1} \sum_{i=1}^{N_1} \left\{ \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^{X, N_2}) - \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^X) \right\} \right\|_p \\ &\leq \frac{1}{N_1} \sum_{i=1}^{N_1} \sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} \left\| \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^{X, N_2}) - \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^X) \right\|_p \end{aligned}$$

As the function \bar{f}_n is k_n -Lipschitz, for all $i \in [\![1, N_1]\!]$,

$$\begin{aligned} & \sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} \left\| \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^{X, N_2}) - \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^X) \right\|_p \\ &= \sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} \mathbb{E}_{\bar{\eta}_0} \left[\left| \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^{X, N_2}) - \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^X) \right|^p \right]^{1/p} \\ &\leq \sup_{n \geq 0} \mathbb{E}_{\bar{\eta}_0} \left[k_n^p \left\| \eta_{\theta_{0:n}^i, n}^{X, N_2} - \eta_{\theta_{0:n}^i, n}^X \right\|_{\mathfrak{F}_n}^p \right]^{1/p} \\ &\leq \sup_{n \geq 0} k_n \mathbb{E}_{\bar{\eta}_0} \left[\left\| \eta_{\theta_{0:n}^i, n}^{X, N_2} - \eta_{\theta_{0:n}^i, n}^X \right\|_{\mathfrak{F}_n}^p \right]^{1/p}. \end{aligned}$$

Set $k \triangleq \sup_n k_n$, then

$$\sup_{n \geq 0} \sup_{\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)} \left\| \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^{X, N_2}) - \bar{f}_n(\theta_n^i, \eta_{\theta_{0:n}^i, n}^X) \right\|_p \leq k \sup_{n \geq 0} \mathbb{E}_{\bar{\eta}_0} \left[\left\| \eta_{\theta_{0:n}^i, n}^{X, N_2} - \eta_{\theta_{0:n}^i, n}^X \right\|_{\mathfrak{F}_n}^p \right]^{1/p}$$

From [[4], Corollary 7.4.5] (p. 249), as one assumes that there exists $m_i \geq 1$ for $\theta_{n,n+m_i}^i \in \mathsf{E}_n^\Theta \times \dots \times \mathsf{E}_{n+m_i}^\Theta$,

$$\sup_{n \geq 0} \mathbb{E}_{\bar{\eta}_0} \left[\left\| \eta_{\theta_{0:n}^i, n}^{X, N_2} - \eta_{\theta_{0:n}^i, n}^X \right\|_{\mathfrak{F}_n}^p \right]^{1/p} \leq \frac{a(p)}{\sqrt{N_2}} \left(I + \frac{m_i p_{\theta^i}(e^{3\alpha_{\theta^i}})}{\varepsilon(M_{\theta^i}^X)^3 \epsilon(G_{\theta^i})^{2m_i}} \right),$$

where $I \triangleq \sup_{n \geq 0} I(\mathfrak{F}_n) < \infty$. One concludes easily. \square

4 Asymptotic analysis of the labeled island particle algorithm

This section deals with the asymptotic behavior of the labeled island particle algorithm. Especially, we focus on the almost sure convergence.

Using Theorem 3.1 obtained in Section 3, one can easily get the almost sure convergence of the double estimator $\tilde{\eta}_n^{N_1}$ toward $\bar{\eta}_n$ under the same assumptions as in Theorem 3.1.

Theorem 4.1. *Under the same assumptions as in Theorem 3.1, for all $n \geq 0$ and for every k_n -Lipschitz function $\bar{f}_n \in \text{Osc}_1(\bar{\mathbb{E}}_n)$, one has*

$$\tilde{\eta}_n^{N_1}(\bar{f}_n) \xrightarrow{a.s.} \bar{\eta}_n(\bar{f}_n), \text{ as } N \rightarrow \infty,$$

with $N = N_1 N_2$ such that $N_1 = N^\alpha$ and $N_2 = N^{1-\alpha}$ for all $\alpha \in]0, 1[$.

Proof. Let $\bar{f}_n \in \text{Osc}_1(\mathbb{E}_n)$ be a k_n -Lipschitz function and $\varepsilon > 0$ a real constant. For all $p \in \mathbb{N}^*$, by Markov's inequality, one has

$$\mathbb{P}(|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)| > \varepsilon) \leq \frac{\mathbb{E}_{\bar{\eta}_0} [|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)|^p]}{\varepsilon^p}.$$

Then, applying Theorem 3.1, and noting

$$C(p, n) \triangleq k_n a(p)(I(\mathcal{F}_n) + b(n)) \quad \text{and} \quad \tilde{C}(p, n) \triangleq 2d(p) \sum_{q=0}^n \bar{g}_{q,n} \beta(\bar{P}_{q,n}),$$

one has

$$\begin{aligned} \|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p^p &\leq \left(\frac{C(p, n)}{\sqrt{N^\alpha}} + \frac{\tilde{C}(p, n)}{\sqrt{N^{1-\alpha}}} \right)^p \\ &= \sum_{k=0}^p \binom{p}{k} \frac{C(p, n)^k \tilde{C}(p, n)^{p-k}}{N^{(\alpha-\frac{1}{2})k + \frac{1-\alpha}{2}p}}. \end{aligned}$$

The finite sequence $(s_{\alpha,p}(k))_{k=0}^p$ defined by $s_{\alpha,p}(k) = (\alpha - 1/2)k + (1 - \alpha)p/2$ is bounded from below by

$$m_{\alpha,p} \triangleq \frac{\alpha p}{2} \mathbf{1}_{0 < \alpha \leq 0.5} + \frac{(1 - \alpha)p}{2} \mathbf{1}_{0.5 < \alpha < 1},$$

so that

$$\|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p^p \leq \frac{\left(C(p, n) + \tilde{C}(p, n) \right)^p}{N^{m_{\alpha,p}}}.$$

Choose p a positive integer such that $m_{\alpha,p} \geq 2$ i.e. satisfying

$$\begin{cases} p > \frac{4}{\alpha} & \text{if } 0 < \alpha \leq 0.5 \\ p > \frac{4}{1-\alpha} & \text{if } 0.5 < \alpha < 1. \end{cases} \quad (25)$$

Hence,

$$\|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)\|_p^p \leq \frac{\left(C(p, n) + \tilde{C}(p, n) \right)^p}{N^2}.$$

By comparison of series of non-negative general term with a convergent Riemann series, one concludes that the series

$$\sum_{N \geq 0} \mathbb{P}(|\tilde{\eta}_n^{N_1}(\bar{f}_n) - \bar{\eta}_n(\bar{f}_n)| > \varepsilon) \text{ is convergent,}$$

which implies by Borel-Cantelli's lemma, that

$$\tilde{\eta}_n^{N_1}(\bar{f}_n) \xrightarrow{a.s.} \bar{\eta}_n(\bar{f}_n), \text{ as } N \rightarrow \infty.$$

□

Taking such kind of N means that for a total budget N of particles, one can consider any decomposition (as a power of N) of the particles between islands and within each island.

5 Example of application

In order to give illustration of this algorithm and of the previous theoretical results obtained, we present in this section two estimation problems.

First let us recall the example of a mobile whose evolution is influenced by an unknown force which has been described in (1). Noisy observations of this physical systems are available. We resume the dynamics by the following system of equations :

$$\begin{cases} X_{n+1} &= X_n + V_n \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \Delta t + \Theta_{n+1} \Delta t + B_n^X \\ V_{n+1} &= V_n + B_n^V \\ Y_n &= h(X_n, V_n) + B_n^Y \end{cases} \quad (26)$$

where X_{n+1} denotes the position of the mobile in the plane, V_n the proper speed of the mobile and Y_n their noisy observations through the observation function h , with $B_n^Y \sim N(0, \Sigma^Y)$. The course track of the mobile α is constant over time. The vector Θ_n is a random variable and denotes the unknown force acting on the position of the mobile. Its equation of evolution is given by

$$\Theta_{n+1} = \begin{pmatrix} \Theta_{n+1}^1 \\ \Theta_{n+1}^2 \end{pmatrix} = \begin{pmatrix} \cos \Theta_n^1 \\ \sin \Theta_n^2 \end{pmatrix} + B_n^\Theta$$

with $B_n^\Theta \sim N(0, \Sigma^\Theta)$. The initial condition of the system is given by $X_0 \sim N(m_{\theta_0, 0}^X, \Sigma_{\theta_0, 0}^X)$, $V_0 \sim N(m_0^V, \Sigma_0^V)$ and $\alpha = \pi/2$. We are interested in the estimation of the position of the mobile, which depends on the parameter Θ_n . We thus need to learn both the force, the speed and the position of the mobile. The tricky part is that there is no observation of the force. Here we will consider that the speed is a Poisson process, that is B_n^V is a Poisson process of intensity 0.03 where the jumps high is given by a standard normal distribution of variance 3. Concerning B_n^X , it is a Gaussian random variable such that $B_n^X \sim N(0, \Sigma^X)$. We present now the results obtained for a simulating time of 125 minutes with $\Delta t = 15s$. The value of the different variances are set to

$$\Sigma^\theta = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \Sigma^X = \Sigma_{\theta_0, 0}^X = \begin{pmatrix} 1.5 & 0 \\ 0 & 1.5 \end{pmatrix}, \text{ and } \Sigma^Y = \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

As one can notice, to estimate the law of the couple $(\Theta_n, \eta_{\Theta_0, n}^X)$ given the observations $Y_{0:n}$, one can use Interacting Kalman filters and labeled island particle filters (LIPFs), detailed respectively in Algorithms 1 and 2. We present comparative results obtained thanks to both methods.

Concerning the labeled version, the potential of each particle is given by the density of the observations, that is for all $x_n \in E_n^X$ and for all $\theta_n \in E_n^\Theta$:

$$G_n(\theta_n, x_n) \propto \exp \left(-\frac{1}{2} (y_n - h(x_n, v_n))^T (\Sigma^Y)^{-1} (y_n - h(x_n, v_n)) \right).$$

On all the figures the realization of the true signal is represented by the color black, the observations Y are represented by the color blue, the filtered signal obtained thanks to Algorithm 2 with $N_1 = 100$ and $N_2 = 300$ in red and results obtained using Algorithm 1 in green with $N_1 = 100$. On Figure 3, one realization of the signal V_n , its observed and its estimations counterparts are represented with respect to time. As one may observe, the true signal is well estimated by the technique we develop. Indeed, here the Interacting Kalman filter is not optimal as the noise sequence is not Gaussian. On Figure 5, we represent the temporal evolution of the force strength estimation. One can notice that even if no observation is available, we are able to find back

the value of the true signal thanks to Algorithm 2 whereas Algorithm 1 retrieves only a global trend. Figure 4 represents the temporal evolution of one realization of the force orientation and its estimated counterparts. Results obtained thanks to Algorithm 2 give a better estimation of the true signal than the results obtained thanks to the Algorithm 1. From this example we can conclude that the labeled island particle filter is able to filter observations of the process while estimating the environment where the process evolves. Moreover the comparison with the Interacting Kalman filter algorithm shows that the labeled island particle filter is more effective to treat this double level estimation problem.

Let us consider the 2-D filtering problem inspired from the growth model [12]. This model, which is a standard benchmark example in the particle filtering literature, is given by the following system of equations :

$$\begin{cases} \Theta_{n+1} &= 8 \cos(1.2(n+1)) + B_{n+1}^\theta \\ X_{n+1} &= \frac{X_n}{2} + 25 \frac{X_n}{1+X_n^2} + \Theta_{n+1} + B_{n+1}^X \\ Y_n &= X_n + B_n^Y \end{cases}$$

where $\Theta_0 \sim N(0, \sigma_\theta^2)$, $X_0 \sim N(0, \sigma_X^2)$, $B_{n+1}^\theta \sim N(0, \sigma_\theta^2)$, $B_{n+1}^X \sim N(0, \sigma_X^2)$ and $B_n^Y \sim N(0, \sigma_Y^2)$. We use the labeled island particle model to estimate the law of the couple $(\Theta_n, \eta_{\Theta_{0:n}, n}^X)$ given the observations $Y_{0:n}$, where the potential functions G_n are given by the likelihood of the observations, that is for all $x_n \in E_n^X$ and $\theta_n \in E_n^\Theta$:

$$G_n(\theta_n, x_n) \propto \exp\left(-\frac{(Y_n - x_n)^2}{2\sigma_Y^2}\right).$$

We present the results obtained for a simulating time of 1000 time steps. The different variances are set to $\sigma_\theta^2 = 1$, $\sigma_X^2 = 1$ and $\sigma_Y^2 = 10$. On all the figures the realization of the true signal is represented in black color, the observations Y are represented in blue, and the filtered signal obtained thanks to Algorithm 2 with $N_1 = 200$ and $N_2 = 100$ is represented in red. On Figure 6, one realization of the signal Θ and its estimation obtained thanks to the labeled island particle algorithm are represented on a small period of time. As one may observe, the true signal is well estimated even if no observations are available. On Figure 7, one realization of the process X is represented, its observed and its estimation counterparts. Even if the observations are really noisy, one is able to filter out the noise to find back the value of the true signal. Indeed, as one may have noticed, on Figure 8, the filtered power spectral density (in red) is closer to the black line, representing the “true” signal, than the observed power spectral density which has the same shape as a white noise for the high frequencies. Moreover, some frequencies are found even if there are not present in the observed signal. These two observations illustrate the convergence of the estimator constructed by the labeled island particle algorithm detailed in Algorithm 2.

Then we run 100 times the same experiment to get a sample of realizations for the true signal and the filtered signal. In that way one can illustrate the theoretical results obtained for the L^p error bound. On figures 9 and 10 are presented the L^2 errors between the estimated law and the true law at one time step respectively for Θ and X in function of the number of islands N_1 and the number of particles inside each island N_2 . This error decreases both with the number of particles and the number of islands as it was suggested by the Theorem 3.1. Concerning the variance of the error made between the true law and the filtered one, on figures 11 and 12 for Θ and X respectively, one can observe that the results obtained in Theorem 4.1 are confirmed. Moreover one can notice that the variance is more influenced by the number of islands than the number of particles inside each island. Indeed as in Figure 12, the variance obtained for a fixed time step is varying with respect to the number of islands and number of particles inside each

islands. But if the number of islands influences the variance, we can observe that the number of particles inside each island does not seem to be really influent for a given number of islands.

References

- [1] R. Cont, “Model uncertainty and its impact on the pricing of derivative instruments,” *Mathematical Finance*, vol. 16, no. 3, pp. 519–547, 2006. [Online]. Available: <http://dx.doi.org/10.1111/j.1467-9965.2006.00281.x>
- [2] J.-C. Augustin and V. Carlier, “Mathematical modelling of the growth rate and lag time for listeria monocytogenes,” *International Journal of Food Microbiology*, vol. 56, no. 1, pp. 29 – 51, 2000. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0168160500002233>
- [3] C. Ichard and C. Baehr, “Inference of a random environment from random process realizations: Formalism and application to trajectory prediction,” in *ISIATM 2013, 2nd International Conference on Interdisciplinary Science for Innovative Air Traffic Management*, 2013.
- [4] P. Del Moral, *Feynman-Kac formulae: genealogical and interacting particle systems with applications*. Series: Probability & Applications Springer Verlag, 2004.
- [5] M. Zghal, L. Mevel, and P. Del Moral, “Modal parameter estimation using interacting kalman filter,” *Mechanical Systems and Signal Processing*, vol. 47, no. 1-2, pp. 139 – 150, 2014, mSSP Special Issue on the Identification of Time Varying Structures and Systems. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327012004633>
- [6] C. Baehr, “Modélisation probabiliste des écoulements atmosphériques turbulents afin d’en filtrer la mesure par approche particulaire,” Ph.D. dissertation, 2008. [Online]. Available: <http://www.theses.fr/2008TOU30113>
- [7] C. Vergé, C. Dubarry, P. Del Moral, and E. Moulines, “On parallel implementation of sequential Monte Carlo methods: the island particle model,” *Statistics and Computing*, vol. 25, no. 2, pp. 243–260, march 2015. [Online]. Available: <http://link.springer.com/article/10.1007%2Fs11222-013-9429-x>
- [8] C. Vergé, P. Del Moral, E. Moulines, and J. Olsson, “Convergence properties of weighted particle islands with application to the double bootstrap algorithm,” 2014, preprint. [Online]. Available: <http://arxiv.org/pdf/1410.4231v1.pdf>
- [9] N. Chopin, P. E. Jacob, and O. Papaspiliopoulos, “SMC²: an efficient algorithm for sequential analysis of state space models,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 75, no. 3, pp. 397–426, 2013.
- [10] N. Chopin, “A sequential particle filter method for static models,” *Biometrika*, vol. 89, no. 3, pp. 539–552, 2002.
- [11] D. Crisan and J. Miguez, “Nested particle filters for online parameter estimation in discrete-time state-space markov models,” 2013, preprint. [Online]. Available: <http://arxiv.org/pdf/1308.1883.pdf>
- [12] G. Kitagawa, “Non-Gaussian state space modeling of nonstationary time series,” *Journal of the American Statistical Association*, vol. 82, no. 400, pp. 1023–1063, December 1987.

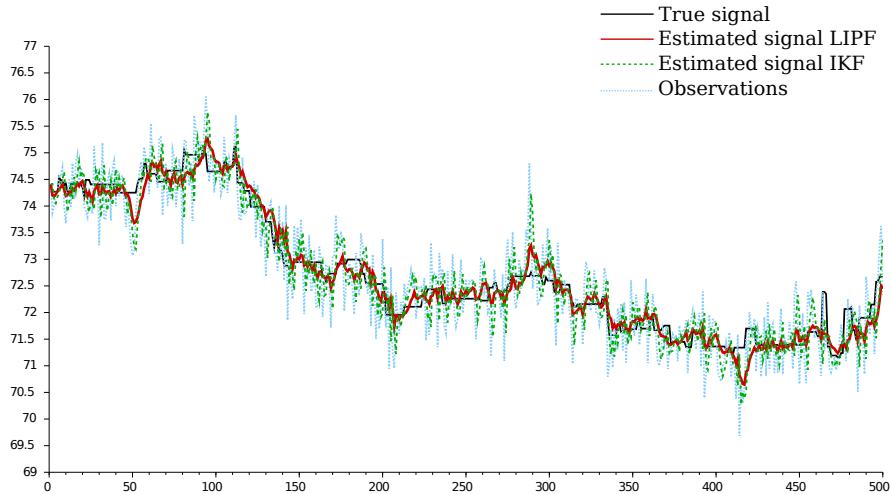


Figure 3: Temporal evolution of the mobile's speed, its observed and filtered counterparts.

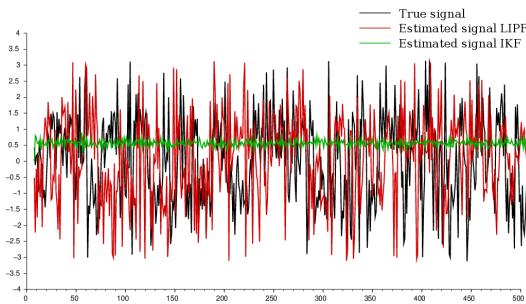


Figure 4: Temporal evolution of the force orientation in rad

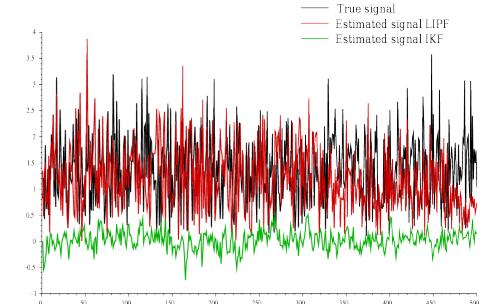


Figure 5: Temporal evolution of the force strength in $m.s^{-1}$

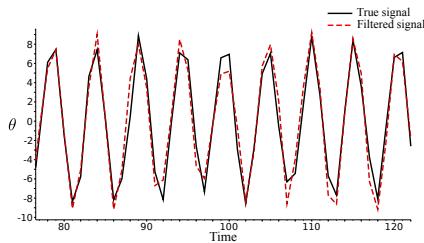


Figure 6: Temporal zoom on one realization of the Θ process and its estimated counterpart

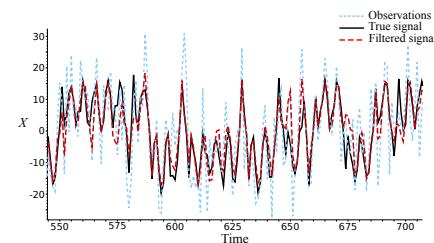


Figure 7: Temporal zoom on one realization of the X process, its observed and estimated counterparts

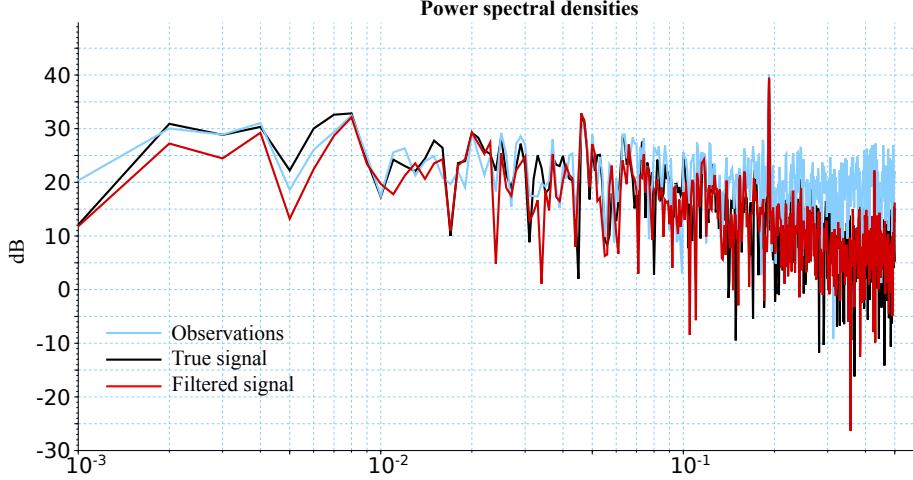


Figure 8: True, filtered and observed power spectral densities for one realization of the process X over 1000 time steps.

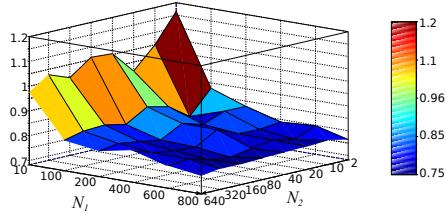


Figure 9: Evolution of the estimation's error for the law of Θ for 100 realizations of the process in function of N_1 and N_2

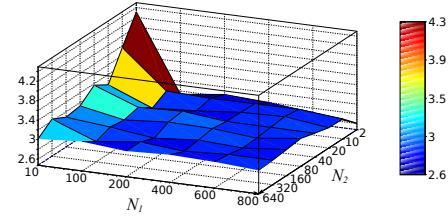


Figure 10: Evolution of the estimation's error for the law of X for 100 realizations of the process in function of N_1 and N_2

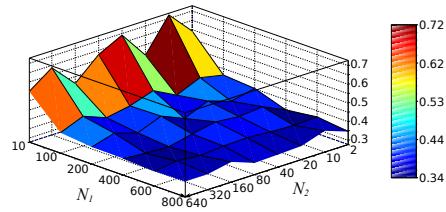


Figure 11: Evolution of the variance estimation's error for the law of Θ for 100 realizations of the process in function of N_1 and N_2

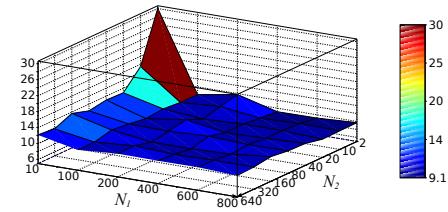


Figure 12: Evolution of the variance estimation's error for the law of X for 100 realizations of the process in function of N_1 and N_2